7-6-2017

Modeling observations of solar coronal mass ejections with heliospheric imagers verified with the Heliophysics System Observatory

C. Mostl
Austrian Academy of Sciences

A. Isavnin
University of Helsinki

P. D. Boakes
Austrian Academy of Sciences

E. K. J. Kilpua
University of Helsinki

J. A. Davies
Rutherford Appleton Laboratory

Recommended Citation

Moestl, C.; Isavnin, A.; Boakes, P. D.; Kilpua, E. K. J.; Davies, J. A.; Harrison, R. A.; Barnes, D.; Krupar, V.; Eastwood, J. P.; Good, S. W.; Forsyth, R. J.; Bothmer, V.; Reiss, M. A.; Amerstorfer, T.; Winslow, R. M.; Anderson, B. J.; Philpott, L. C.; Rodriguez, L.; Rouillard, A. P.; Gallagher, P.; Nieves-Chinchilla, T.; Zhang, T. L. (2017). Modeling observations of solar coronal mass ejections with heliospheric imagers verified with the Heliophysics System Observatory, SPACE WEATHER-THE INTERNATIONAL JOURNAL OF RESEARCH AND APPLICATIONS. Vol. 15, No. 7, 955-970. 10.1002/2017SW001614

AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2017SW001614

Key Points:
- Hindsight predictions of CMEs by modeling observations from the STEREO heliospheric imagers science data are analyzed for 2007-2014
- About one out of four predicted CME arrivals is observed in situ as a magnetic obstacle by MESSENGER, Venus Express, Wind, and STEREO A/B
- Although using strong assumptions on CME physics, prediction accuracies for CME arrival times are similar to numerical or analytical models

Correspondence to:
C. Möstl,
christian.moestl@oeaw.ac.at

Citation:
Möstl, C., et al. (2017), Modeling observations of solar coronal mass ejections with heliospheric imagers verified with the Heliophysics System Observatory, *Space Weather*, *15*, 955–970, doi:10.1002/2017SW001614.

Received 22 FEB 2017
Accepted 30 JUN 2017
Accepted article online 6 JUL 2017
Published online 29 JUL 2017

Modeling observations of solar coronal mass ejections with heliospheric imagers verified with the Heliophysics System Observatory

C. Möstl[1,2], A. Isavnin[3], P. D. Boakes[1,2], E. K. J. Kilpua[3], J. A. Davies[4], R. A. Harrison[4], D. Barnes[4,5], V. Krupar[6], J. P. Eastwood[7], S. W. Good[7], R. J. Forsyth[7], V. Bothmer[8], M. A. Reiss[2], T. Amerstorfer[1], R. M. Winslow[9], B. J. Anderson[10], L. C. Philpott[11], L. Rodriguez[12], A. P. Rouillard[13,14], P. Gallagher[15], T. Nieves-Chinchilla[16], and T. L. Zhang[1]

[1]Space Research Institute, Austrian Academy of Sciences, Graz, Austria, [2]IGAM-Kanzelhöhe Observatory, Institute of Physics, University of Graz, Graz, Austria, [3]Department of Physics, University of Helsinki, Helsinki, Finland, [4]RAL Space, Rutherford Appleton Laboratory, Harwell, UK, [5]University College London, London, UK, [6]Institute of Atmospheric Physics CAS, Prague, Czech Republic, [7]Blackett Laboratory, Imperial College London, London, UK, [8]Institute for Astrophysics, University of Göttingen, Göttingen, Germany, [9]Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA, [10]Applied Physics Laboratory, The Johns Hopkins University, Laurel, Maryland, USA, [11]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, British Columbia, Canada, [12]Solar Terrestrial Center of Excellence-SIDC, Royal Observatory of Belgium, Brussels, Belgium, [13]Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse (UPS), Toulouse, France, [14]Centre National de la Recherche Scientifique, Toulouse, France, [15]School of Physics, Trinity College Dublin, Ireland, [16]Heliophysics Science Division, GSFC/NASA, Greenbelt, Maryland, USA

Abstract We present an advance toward accurately predicting the arrivals of coronal mass ejections (CMEs) at the terrestrial planets, including Earth. For the first time, we are able to assess a CME prediction model using data over two thirds of a solar cycle of observations with the Heliophysics System Observatory. We validate modeling results of 1337 CMEs observed with the Solar Terrestrial Relations Observatory (STEREO) heliospheric imagers (HI) (science data) from 8 years of observations by five in situ observing spacecraft. We use the self-similar expansion model for CME fronts assuming 60° longitudinal width, constant speed, and constant propagation direction. With these assumptions we find that 23%–35% of all CMEs that were predicted to hit a certain spacecraft lead to clear in situ signatures, so that for one correct prediction, two to three false alarms would have been issued. In addition, we find that the prediction accuracy does not degrade with the HI longitudinal separation from Earth. Predicted arrival times are on average within 2.6 ± 16.6 h difference of the in situ arrival time, similar to analytical and numerical modeling, and a true skill statistic of 0.21. We also discuss various factors that may improve the accuracy of space weather forecasting using wide-angle heliospheric imager observations. These results form a first-order approximated baseline of the prediction accuracy that is possible with HI and other methods used for data by an operational space weather mission at the Sun-Earth L5 point.

Plain Language Summary Solar storms are formed by incredibly powerful explosions on the Sun and travel as clouds of plasma threaded by magnetic fields through the solar system. Depending on their propagation direction, they may impact planets such as Earth, where they elicit colorful aurorae or, in very seldom cases, can lead to power failures with potentially tremendous economical and societal effects, thus posing a serious natural hazard. In this work, we have shown how well the solar storm impact can be forecasted when using a special type of instrument that can actually image the solar storms as they propagate toward the planets and even as they sweep over them. Our analysis includes two thirds of a solar cycle with 8 years of data, and spacecraft at Mercury, Venus, Earth, and in the solar wind to check on the correctness of our predictions. We could forecast the arrival time within ±16 h, and for one correct impact there are two to three false alarms. This forms a new baseline for the science of space weather prediction. Clearly, the modeling should be further improved to be used on a daily basis for a space weather mission to the Sun-Earth L5 point.

1. Introduction

A major goal in space weather research is to improve the accuracy of forecasts concerning coronal mass ejection (CME) arrival times and speeds, and whether a CME impacts or misses the Earth. This is needed to ensure that potentially disrupting effects for society are mitigated [*Oughton et al*., 2017, and references therein]. The heliospheric imagers (HI [*Howard et al*., 2008; *Eyles et al*., 2009]) on board Solar Terrestrial Relations Observatory (STEREO) have pioneered this approach for the first time away from the Sun-Earth line. Because solar wind structures such as CMEs can be observed as they propagate toward the Earth and even as they sweep over it [*Davis et al*. 2009; *Möstl et al*. 2010; *Liu et al*. 2010; *Colaninno et al*. 2013; *Möstl et al*. 2014; *DeForest et al*., 2016; *Wood et al*. 2017], it makes sense to use modeling approaches specifically designed for HI [*Liu et al*., 2010; *Lugaz*, 2010; *Davies et al*., 2012; *Möstl and Davies*, 2013; *Möstl et al*., 2014; *Rollett et al*., 2016] to predict CME arrival times and whether a CME will impact a certain heliospheric location, similar to analytical or numerical modeling [e.g., *Vršnak et al*., 2014; *Mays et al*., 2015]. For assessing the prediction accuracy, data not only of the near Earth solar wind but also from several other spacecraft operating in the heliosphere can be used. This collection of spacecraft is known as the Heliophysics System Observatory (HSO). Here we use data from the HSO spacecraft operating in the heliosphere and at the Sun-Earth L1 point, but no data from other spacecraft in geospace.

In this study, we test the validity of using HI observations for space weather forecasting, in particular for the CME arrival time, and the accuracy of hit and miss predictions. These are essential results for a possible future space weather mission to the L5 point, which could continuously monitor the Sun and the space between the Sun and Earth with HI [*Lavraud et al*., 2016; *DeForest et al*., 2016]. Wide field imagers, somewhat akin to the STEREO/HI instrument, will also be onboard Solar Orbiter and Solar Probe Plus, and STEREO Ahead/HI continues to observe the solar wind between the Sun and Earth in July 2015, emphasizing that our results will provide a baseline for future missions.

We present here the first work in which predictions based on a large-scale sample of 1337 CMEs observed with either HI instrument (HIA on STEREO A and HIB on STEREO B) are tested with 641 interplanetary coronal mass ejection events observed in situ, obtained between April 2007 and September 2014 and ranging in heliocentric distances from 0.308 to 1.086 AU. It has been shown previously [e.g., *Colaninno et al*., 2013; *Möstl et al*., 2014; *Tucker-Hood et al*., 2015] that STEREO/HI may enhance the warning times and the accuracy of CME predictions, even in near real time [*Tucker-Hood et al*., 2015]. However, the conclusion of these studies was that there is still room for improvement concerning the HI modeling and that the maximum number of events that could be ratified with a corresponding in situ observation was around 20 [e.g., *Möstl et al*., 2014; *Tucker-Hood et al*., 2015]. Note that in this present work, all predictions are in fact made in hindsight.

For predicting space weather at Earth or any other planet it is necessary to test the HI modeling techniques also for distances other than 1 AU. Studies so far have focused on in situ data near 1 AU [*Möstl et al*., 2014; *Rollett et al*., 2016] or on data taken in the near-Earth solar wind [*Tucker-Hood et al*., 2015]. Our aim in this work is thus twofold: (1) to substantially increase the number of CME events available for the statistical prediction performance analyses and (2) to test the forecasts based on HI data for heliocentric distances other than 1 AU. This is of relevance for studies of planetary space weather and even for manned Mars missions.

2. Methods and Data

This work is an outcome of the *HELiospheric Cataloguing*, *Analyses*, *and Techniques Service* (HELCATS) project, in which eight academic research institutions from around Europe were involved from May 2014 to April 2017. During the time of this project, we have established, among many other catalogs from heliospheric data sets, several catalogs that are relevant for this study: the catalog of observed CMEs with HI is called HIGeoCAT, the catalog of predicted arrivals ARRCAT, the catalog of the in situ interplanetary data DATAC AT, and information on the in situ CME (ICME) observations is gathered in the ICMECAT. It contains data from the solar wind observatories Wind, STEREO A, STEREO B, Venus Express (VEX), MErcury Surface, Space ENvironment, GEochemistry, and Ranging (MESSENGER), and Ulysses. We use all spacecraft for a comparison to HI except Ulysses because there are too few CME events observed by this spacecraft.

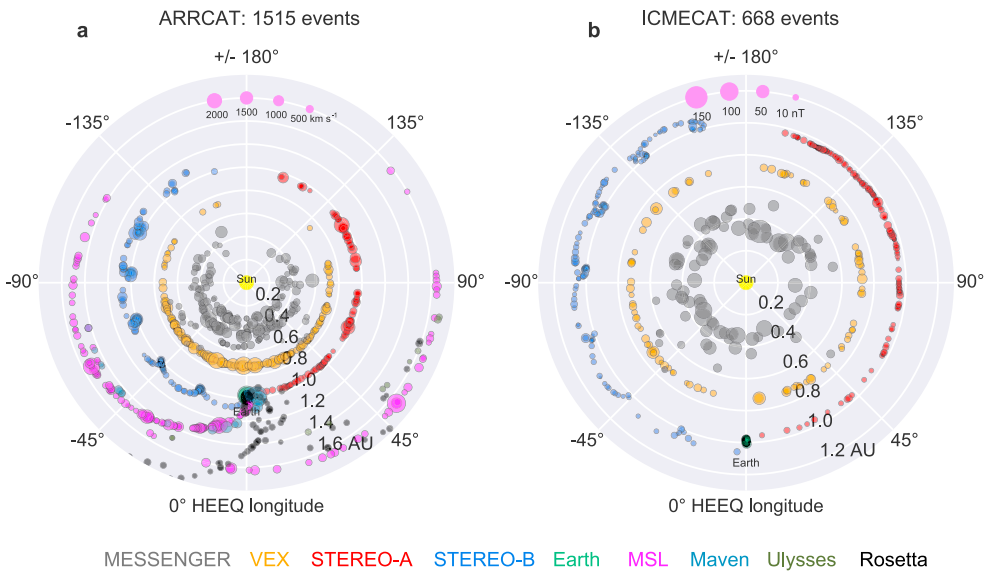

Figure 1. Overview of ARRCAT and ICMECAT. Both panels show the solar equatorial plane. (a) All CMEs in HIGeoCAT were checked if they potentially arrive at various planets and spacecraft as indicated by the color code at the bottom, based on the shape model of a self-similar expanding circle with 30° half width (SSEF30). Each dot marks a predicted arrival at MESSENGER, VEX, STEREO A/B, Earth/L1, MSL, MAVEN, and Ulysses and Rosetta. The size of the dot indicates the predicted impact speed, which is an overestimate due to the constant speed assumption of SSEF30. (b) Overview of in situ detections of ICMEs, showing the longitude and radial distance at which the detection happened, as collected in ICMECAT. The size of the circle indicates the mean magnetic field strength in the magnetic obstacle. Here MESSENGER, VEX, STEREO A/B, and Earth/L1 are shown.

2.1. HIGeoCAT

The CME catalog, HIGeoCAT, forms the basis of our investigation. This is underpinned by a catalog of CMEs visually identified in images from the inner (HI1) camera on each STEREO spacecraft (available on the HELCATS website, see section 5), using science data (not real-time beacon data). For each CME in that catalog categorized as either "good" or "fair," a time-elongation map (commonly called a J map) was constructed from combined HI1 and HI2 difference observations, along a position angle corresponding to—as near as practicable—the CME apex. Note that in 71% of cases, this is within 30° of the solar equatorial plane, which is situated at a position angle of 90 (HIA) and 270° (HIB). This demonstrates the tendency for CMEs to be confined to low equatorial latitudes. The time-elongation profile corresponding to the leading edge of the CME was extracted, by clicking along the appropriate trace in the J map [*Sheeley et al.*, 1999; *Rouillard et al.*, 2008]; for each CME, this process was repeated multiple times. The resulting tracks were fitted with the Self-Similar Expansion Fitting (SSEF) method [*Davies et al.*, 2012] with 30° half width, in short called SSEF30. Thus, for each CME a 60° full width in the solar equatorial plane is assumed. The method assumes the CME front to be of circular shape and is based on geometrical aspects of objects that propagate at a given angle to the observer. We use only single-spacecraft HI results, similar to what would be available on an L5 mission, and no stereoscopic results are used. The CME parameters in HIGeoCat include among others a constant speed and constant propagation direction, and a launch time *sse_launch*, which is a backprojection to the time of the CME's inception in the solar corona. It is usually very close in time to the first appearance of the CME in the STEREO coronagraph COR2 [*Möstl et al.*, 2014]. The time range of the HI observations and modeling was 1 April 2007 to 27 September 2014.

2.2. ARRCAT

Figure 1a shows an overview of the positions of the predicted CME impacts in the heliosphere. This is based on the ARRival CATalog (ARRCAT) that was produced for all events in HIGeoCAT. ARRCAT contains impacts at Earth L1, STEREO A, STEREO B, VEX, and MESSENGER. These are further called the targets. We have also produced arrival lists for Mars, Saturn, Ulysses, Mars Science Laboratory (MSL), MAVEN, and Rosetta, but they are not used in this work. The position of VEX is assumed equal to the location of Venus, and the location of MESSENGER is equal to Mercury after orbit insertion on 18 March 2011.

Essentially, for each CME it is assessed whether a planet or spacecraft is currently situated within ±30° heliospheric longitude (Heliocentric Earth Equatorial, HEEQ [*Thompson*, 2006]) of the CME propagation direction. If this is the case, the SSEF30 circle is expected to impact a target, and an entry in ARRCAT is produced. The heliospheric positions of those targets are assessed at the time *sse_launch* as the CME leaves the Sun. For creating a predicted arrival time of a CME, we add a time of travel, depending on the heliocentric distance of the target and the CME speed, to *sse_launch* (for details on the method see *Möstl and Davies* [2013]. The predicted impact time *target_arrival* and the speed are calculated including the circular shaped SSEF30 front, which results in significant delays of flank impacts compared to head-on hits due to the circular geometry [*Möstl and Davies*, 2013].

At VEX and MESSENGER, the spacecraft position at *sse_launch* is actually a few degrees away from the planet position at the time of the hit. For STEREO A/B and Earth, this problem does not occur because in the HEEQ coordinate system, their positions move very little (STEREO) or not at all (Earth). Therefore, a slight shift occurs between the longitudinal position of the predicted impact and the actual longitude of the targets VEX and MESSENGER at impact time. For VEX, this amounts on average to 1.8 ± 0.9° longitude, and for MESSENGER its 4.2 ± 2.6°.

It is not straightforward to calculate how this shift affects the arrival time calculation, because this depends on many factors such as CME speed, heliocentric distance, and in particular the position of the impact along the circular front. However, we can make an estimate based on average values. With the method from *Möstl and Davies* [2013] we find that for an average predicted impact speed of 513 km s^{-1} at the distance of Venus, and a difference to the CME apex of 15° the difference in arrival time due to a shift of 1.8° is about 1.3 h. For very extreme cases (very slow CME, far flank hits) it can also be on the order of 10 h, which is caused by the strong curvature of the circular SSEF shape at the front edges. A similar value of 1.7 h difference in arrival time is found for MESSENGER, again for average values. This means that the arrival time calculation is very often not much affected by the systematic error in assessing the spacecraft position at launch time. Nevertheless, this is a problem inherent to our current method of arrival time calculation and should be fixed in future updates.

2.3. DATACAT

The name DATACAT is given to the collection of all files that include the in situ magnetic field and plasma data, downloaded from various online sources at the respective sites of the missions, and converted to a single coordinate system. The instrument studies of the magnetometers are for Wind *Lepping et al*. [1995], for VEX *Zhang et al*. [2006], for STEREO *Luhmann et al*. [2008], and for MESSENGER *Anderson et al*. [2007]. For VEX and MESSENGER, only magnetic field data are available. For completeness, plasma speeds, temperatures, and densities were taken from STEREO (PLASTIC [*Galvin et al*., 2008]) and Wind (SWE [*Ogilvie et al*., 1995]). Except for a comparison to the plasma speeds at Wind, we do not use the plasma data in this study. All data were either available or resampled to a 1 min time resolution.

The in situ data start on 1 January 2007. MESSENGER was still in cruise phase and has a few long intervals when no data was taken. The magnetic field data from the other missions are almost continuous. The magnetosphere measurements have been removed from MESSENGER observations in Mercury orbit, starting with the orbit insertion at Mercury on 18 March 2011, based on a manually established list of magnetosphere crossings for every MESSENGER orbit [*Winslow et al*., 2013] (updated until the end of the mission). For each orbit, we have excluded all data that were taken in between the outer boundaries of the bow shock. We have also manually removed any residual spikes (defined by clear instrumental artifacts of singular measurements above 100 nT) in the magnetic field data, so the calculations of ICME parameters are not affected. Calibrated level 2 magnetic field and plasma data from STEREO In situ Measurements of Particles And CME Transients (IMPACT) and Plasma and Suprathermal Ion Composition (PLASTIC) instruments, respectively, were obtained from the Space Physics Center of University of California, Los Angeles. Wind magnetic field and plasma data from MFI and SWE instruments were obtained from Coordinated Data Analysis Web service maintained by NASA. For removing the induced magnetosphere of Venus from the VEX data, a slightly modified formula from *Zhang et al*. [2008] was used. We use a terminator crossing of 3.5 instead of 2.14 and 2.364 [*Zhang et al*., 2008], which works to exclude almost all movements of the bow shock while still retaining large amounts of solar wind data.

Due to the end of the missions, VEX data terminates on 26 November 2014, and data from MESSENGER on 30 April 2015. STEREO A moved to solar conjunction starting with 19 August 2014, and contact with STEREO B was lost on 27 September 2014. Wind data are available still after than 27 September 2014, but the data used in our study end at this time because no continuous in situ nor imaging data from either STEREO spacecraft are available from this date to July 2015, when STEREO A continued observations.

2.4. ICMECAT

We define an interplanetary coronal mass ejection (ICME) as the full interval of disturbed solar wind, including the shock, sheath, magnetic obstacle, and a possible wake region [*Rouillard*, 2011; *Kilpua et al*., 2013]. We have gathered lists of in situ CME observations from various online catalogs and published sources, in the time frame January 2007 to December 2015, which overlaps fully with the HIGeoCAT. This results in 668 ICME events. For the comparison with HI, from April 2007 to September 2014, 641 events at the five target spacecraft are available. The reduction from 668 events to 641 is caused by the time ranges January 2007 to March 2007 and October 2014 to December 2015 not having continuous HI observations and because the few events observed by Ulysses have been omitted for comparison to HI.

Figure 1b provides an overview of the positions where those ICMEs have been detected with respect to Earth, demonstrating the 360° heliospheric coverage of ICMEs. It is clear that merging the individual catalogs leads to some problems concerning different definitions of parameters. To work around this, we have taken *only times* from the individual catalogs: from the Wind ICME catalog for L1 ICMEs (by Teresa Nieves-Chinchilla), from the STEREO A/B ICME list by Lan Jian, and from the VEX ICME list from *Good and Forsyth* [2016]. ICMEs at MESSENGER were taken from *Winslow et al*. [2015] and *Good and Forsyth* [2016]. We have updated all ICME lists except for Wind to include ICMEs until the end of the missions (VEX and MESSENGER) or until solar conjunction was reached (STEREO A) or contact was lost (STEREO B).

There are four times included in ICMECAT, for each event. First, there is the *icme_start_time*, which is equal to the shock arrival or the beginning of a significant density enhancement. Then, *mo_start_time* and *mo_end_time* are given which signify the beginning and end of the magnetic obstacle, which is taken as an umbrella term for any significant magnetic structure with higher total magnetic field, and it may include smoothly rotating or complex magnetic field components. If there is no sheath region, i.e., the ICME starts immediately with a magnetic obstacle, the *icme_start_time* is similar to *mo_start_time*. For Wind only, a separate *icme_end_time* is given, for all other spacecraft the *icme_end_time* is the same as the *mo_end_time*. Interacting ICMEs are treated usually as a single ICME in all catalogs, and only if they can be clearly separated they are considered as two individual structures.

ICMEs were identified at STEREO by Lan Jian using a range of plasma and magnetic field signatures, namely, an enhanced total perpendicular pressure (including the contribution of protons, electrons and alpha particles), a declining speed profile, low proton temperature, a relatively enhanced magnetic field strength, and smooth magnetic field rotations [*Jian et al*., 2006; *Jian et al*., 2013]. The presence of at least three of these signatures was required for an ICME identification. A minimum time duration for the signatures was not imposed, although all of the cataloged ICMEs had magnetic obstacle durations of at least 2.5 h.

The magnetic obstacles of ICMEs at Wind were identified as regions of magnetized plasma where the magnetic pressure was significantly greater than the plasma pressure (i.e., where the plasma β was low). Obstacles that showed rotation in at least one field component were defined as flux ropes, and obstacles that showed no rotation were defined as ejecta. Similarly to the STEREO catalog, signature duration was not used as an identification criterion; the shortest magnetic obstacle duration of an ICME observed at Wind was 5.5 h.

For MESSENGER and VEX, only magnetic field data were available for identification of ICMEs and subsequent derivation of the parameters; thus, they need some special attention. ICMEs observed at MESSENGER and VEX that have been listed by *Good and Forsyth* [2016] were all identified by the presence of magnetic flux rope signatures, i.e., relatively smooth rotations of the magnetic field direction coinciding with an elevated magnetic field strength. In general, flux ropes that had a duration of less than 4 h were not included in their catalog. No shock-driving identification criterion was imposed. MESSENGER observed ICMEs during both its long cruise phase and subsequent orbital phase at Mercury (from 18 March 2011); all ICMEs identified at VEX were observed during the spacecraft's orbital phase at Venus. ICMEs observed by MESSENGER and VEX during the orbital phases were obscured to varying degrees by magnetospheric intervals, although ICME boundaries

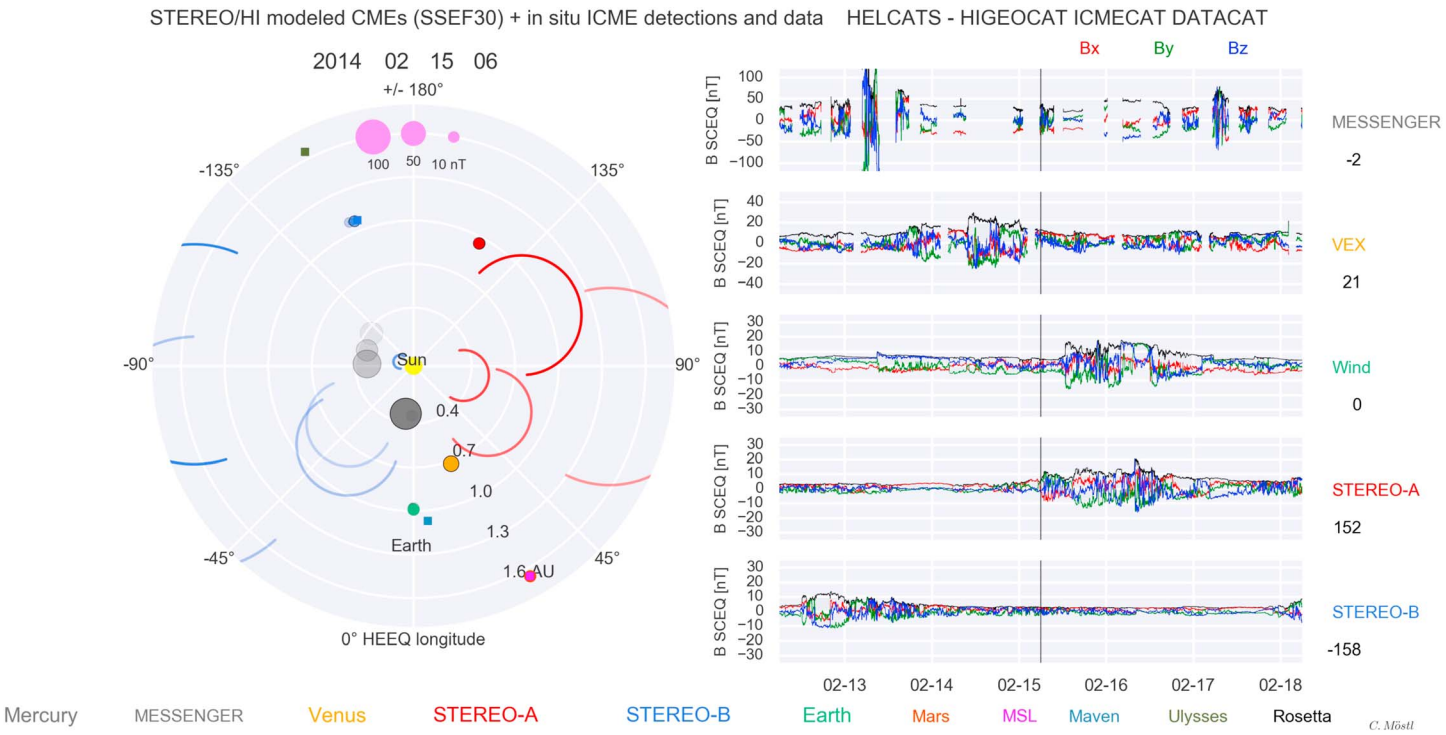

Figure 2. Screenshot of a visualization of CMEs observed with STEREO/HI and their in situ detection by various spacecraft, covering 2007–2015 at a 6 h time resolution. (left) Semicircles denote CMEs that propagate away from the Sun as modeled with SSEF30. Circles that light up and fade around spacecraft positions are actual in situ ICME detections. The size of the circle that lights up is related to the mean magnetic field in the ICME, given by the legend on the upper part. (right) The in situ magnetic field components in colors (B_x red, B_y green, and B_z blue) and the total field (black) are shown from top to bottom for the five spacecraft MESSENGER, VEX, Wind, STEREO A, and STEREO B. The number below each spacecraft label on the far right is the HEEQ longitude at the current animation time, which is shown above Figure 2 (left) and also given by the vertical solid line on each plot in Figure 2 (right).

were clearly visible in most cases. In case of overlaps of the MESSENGER ICME lists by *Winslow et al.* [2015] and *Good and Forsyth* [2016] we have taken the times from *Winslow et al.* [2015] because they include both the ICME start time and the magnetic obstacle times, whereas in the *Good and Forsyth* [2016] list only the magnetic obstacle times were available. We have also added shock times to the list by *Good and Forsyth* [2016] for those events from VEX and MESSENGER where a clear shock preceded the magnetic obstacle.

For all ICMEs, we have derived 30 parameters (e.g., mean and maximum magnetic field, B_z field parameters, and minimum variance analysis of the magnetic obstacle) directly from the data in a homogeneous way, thus eliminating the need to compile different parameters from the different catalogs which might differ in their definitions from one catalog to another. In summary, in ICMECAT 483 ICMEs were observed close to 1 AU (Wind, STEREO A, and STEREO B), and 180 events in the inner heliosphere by VEX and MESSENGER, and five events by Ulysses. However, for our analysis only the *icme_start_time*, the mean magnetic field in the magnetic obstacle and the sheath speed at Wind is used for a comparison to the HI predictions. The catalog is thus introduced here also as a basis for future studies.

2.5. Summary of Data and Methods

Figure 2 is a screenshot of an animation that provides a convenient overview of the spacecraft positions, the CMEs in HIGeoCAT, and the magnetic field components and total field for the full duration of the data we study. The movie has a 6 h time resolution and covers 2007–2015 (see links to the online animation at the end of the paper). In Figure 2 (left), the SSEF30 circles propagate away from the Sun, for the inner heliosphere up to Mars orbit, for all CMEs observed by HI. For each CME, kinematics have been established by calculating $R(t) = V\_sse \times t$, where $R(t)$ is the distance to Sun center, V_sse is the speed from SSEF30, and t is the time, starting with the SSEF launch time. These kinematics are then interpolated to the movie frame times. All apex positions for each CME are then sorted by time. When plotting the movie, for each movie frame, it is checked how many CMEs are currently "active" at that particular moment in time in the spatial domain covering

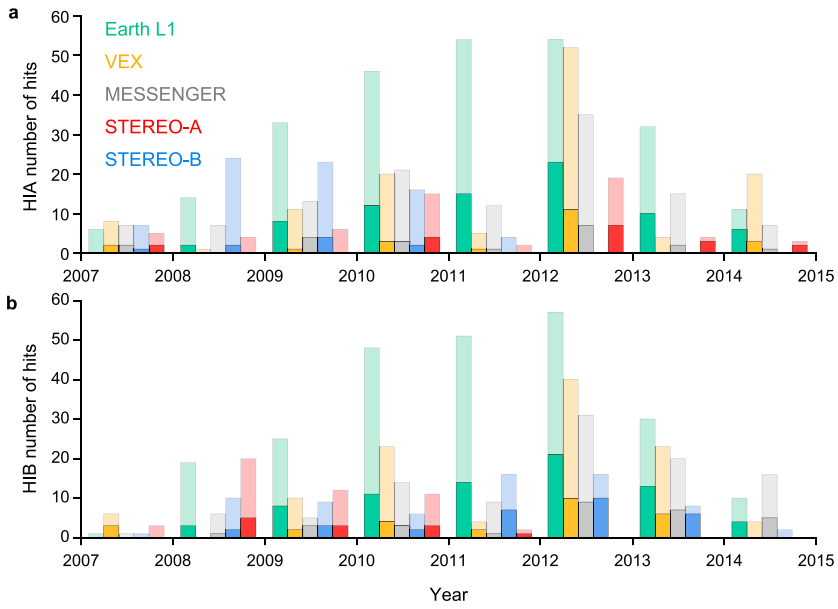

Figure 3. Comparison of correct predictions and false alarms. (a) The number of predicted CME hits with SSEF30 based on HIGeoCAT for each year (only HIA) is shown as shaded bars as function of time. The solid bars are the number of impacts that are accompanied by an actual ICME in situ detection at the respective planet or spacecraft indicated by the color code. The visible shaded part of the bars indicates the number of false alarms. (b) Same based on HIB observations.

approximately 0 to 2.1 AU. The SSE circle is then plotted for each CME with ±110° around the apex position, resulting in a circular arc-like CME front appearance.

Every circle in the movie has a full width of 60°, which was chosen to be consistent with the average CME width in SOHO coronagraph observations [*Yashiro et al.*, 2004]. We know from previous studies [e.g., *Möstl et al.*, 2014] that the accuracy of the CME direction from SSEF30 has some problems—sometimes one can recognize two circles (one red and one blue) which are obviously the same CME seen from STEREO A and B because of the close timing, but the directions differ sometimes on the order of up to 50 degree in extreme cases although they describe the same CME. In Figure 2 (right) we show magnetometer observations by the five spacecraft as collected in DATACAT.

In summary, we have established the catalog of CME arrivals that we call ARRCAT, which is based upon the SSEF30 modeling in HIGeoCAT. We have also gathered in situ observations from up to 9 years of five different spacecraft in the DATACAT, from which we have derived ICME parameters based on the timings in existing ICME lists from the individual spacecraft, which we have updated and merged together in the ICMECAT.

3. Results

To test the predictive capabilities of data from STEREO/HI, we have essentially verified the catalog of predicted arrivals (ARRCAT) with respect to the *icme_start_time* in the catalog of in situ ICME detections (ICMECAT). This shows some fundamental results which are important for space weather forecasting with an HI instrument, based on modeling with the SSEF30 technique.

3.1. Hits and False Alarms

In Figure 3 we assess hit and false alarm predictions: Figure 3a shows the number of predicted CME hits, only from HIA, for each year as shaded bars. Figure 3b shows the same based on HIB observations.

The solid bars represent the number of impacts that are accompanied by an actual ICME in situ detection at the respective planet or spacecraft indicated by the color code. To this end, a time window of ±1.5 days was used around the predicted arrival time to search for ICME start times. Table 1 shows the numbers for correctly predicted impacts at each spacecraft and for HIA/HIB separately. The average for MESSENGER and VEX at distances <1 AU is 22%, for the other spacecraft close to 1 AU it is 25% without self-predictions. With self-predictions we denote CMEs that were observed by the same STEREO spacecraft with HI and in situ. The

Table 1. Percentage of Correct Hits of the ARRCAT Predictions, Meaning There is an Entry in ICMECAT Within ±1.5 Days of a Predicted Arrival in ARRCAT and the Mean and Standard Deviation of Calculated Minus Observed (C – O) Arrival Times, in Hours

Spacecraft	Wind	STEREO A	STEREO B	VEX	MESSENGER	Average
HIA hits, %	30	31[a]	12	17	17	**23**
HIB hits, %	31	25	44[a]	25	28	**30**
HIA C-O, hours	3.2 ± 16.3	1.0 ± 23.4[a]	8.0 ± 14.5	−2.1 ± 11.9	3.2 ± 18.1	**2.4 ± 17.1**
HIB C-O, hours	0.8 ± 17.4	6.8 ± 16.6	7.7 ± 14.7[a]	−0.7 ± 13.7	3.8 ± 12.9	**2.7 ± 16.0**

[a]self-predictions.

accuracy percentage of the self-predictions are significantly higher with 38%, when averaging the numbers from both spacecraft. This underpins the idea that heliospheric imaging might work well also from L1 or Earth orbit [*DeForest and Howard*, 2015; *DeForest et al*., 2016].

Overall, the prediction accuracy is 26%, so one out of four predicted impacts actually caused a clear ICME observed in the solar wind. This compares well to *Tucker-Hood et al*. [2015] who found that 1 out of 3, or 20 predicted arrivals out of 60 in total, lead to an ICME observed at Earth/L1, though their time window was larger. If we choose a different time window than 1.5 days, a window of 1.0 days leads for Earth to a percentage of 24%, 0.5 days to 15%, and 2.0 days to 37%, so the true positives range roughly between one out of three and one out of six, depending which time window is deemed appropriate.

For a time window of 1.5 days, this means that for every correct prediction (solid bars in Figure 3) there are about three false alarms (shaded bars). The average for HIB (30%) is higher than for HIA (23%), which might imply that the "view" from solar east works better compared to solar west, but it is hard to tell if this difference is significant without further analysis. Our study contains a large number of events, i.e., 250 predicted arrivals for Wind could be compared to potential matches within a list of 165 ICMEs, for VEX 121 predictions against 93 ICMEs and for MESSENGER 117 forecasted impacts were compared to 87 ICMEs.

Figure 4 demonstrates how the percentage of correct hits developed (a) as a function of time and (b) separation from Earth in heliospheric longitude. Here only the percentages of correct hits at Earth/L1 are shown, for both HIA and HIB. STEREO B reached the L5 point at 60° separation from Earth in October 2009, and the percentage of correct hits during 2009 was better compared to 2008, where the mean separation was about 30° heliospheric longitude. Surprisingly, the percentage did not decrease with angles larger than L5 but slightly

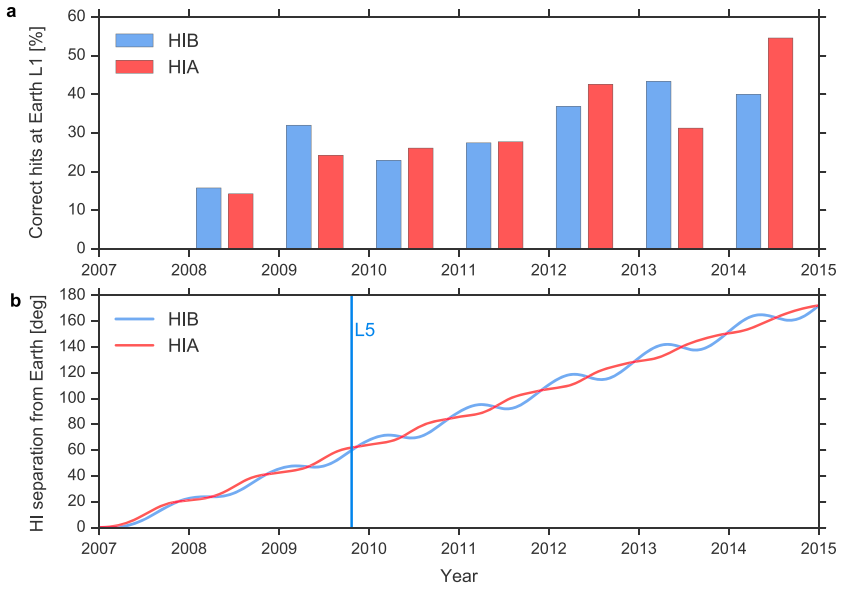

Figure 4. Correct hit percentage as function of time and longitudinal separation to Earth. (a) Yearly percentage of correct hits as function of time. (b) Separation of STEREO A/B from Earth in heliospheric (HEEQ) longitude with time. The moment as STEREO B passed L5 is highlighted as a vertical line.

Table 2. Contingency Table for Earth/L1 for T_{window} = 3 Days[a]

Predicted/Observed	In Situ ICME	No In Situ ICME
ICME arrival	hit (TP): 76 (74)	false alarm (FP): 174 (167)
No ICME arrival	false rejection (FN): 98 (97)	correct rejection (TN): 551 (574)

[a]Numbers Are Given for HIA (HIB).

increased as STEREO A went behind the limb as seen from Earth and finally into conjunction. *Lugaz et al.* [2012] first described the possibility that using HI works for CMEs that propagate behind the limb as seen by an HI spacecraft, and this is consistent with our findings. However, the effect that the percentage of correct hits slightly increases with longitudinal separation needs to be investigated further. The current results imply that L5 is not a particular outstanding point for predicting CME hits and false alarms, but of course still very desirable because of its relative stability.

3.2. Contingency Table Results

The common approach in the event-based validation of space weather forecasts is the use of the contingency table [e.g., for CMEs [*Zhao and Dryer*, 2014; *Tucker-Hood et al.*, 2015], for flares [*Bloomfield et al.*, 2012; *Murray et al.*, 2017], and for high-speed streams [*Reiss et al.*, 2016]). The performance information contained in the contingency table can be presented in form of verification measures, with the ultimate goal to objectively select the best performing method for operational forecasts. There are some precautions to consider, though. First, the entries in the contingency table need to be clearly defined. Second, different verification measures put different emphases on the table entries and may depend on the event/nonevent ratio. Third, a comparison of different methods for CME prediction is only meaningful if the same time range of observations and the same list for in situ observed interplanetary CMEs (ICMEs) is used for CME forecast verification. We would like to emphasize that the catalogs presented here, in particular the ICMECAT containing events at various planets and spacecraft, would be a suitable basis for such investigations with other methods for CME prediction than heliospheric imaging (see section 5 at the end of this study).

With this in mind, in Table 2 we present the 2 × 2 contingency table for Earth/L1. It contains the numbers of hits (true positives, *TP*) and false alarms (false positives, *FP*), and the values for false rejections (false negatives, *FN*, also called "misses") and correct rejections (true negatives, *TN*). TPs are defined by predicted arrivals that are accompanied by an ICME within ±1.5 days of the predicted arrival, FPs are predictions that did not result in an ICME (within ±1.5 days of the predicted arrival), and FNs are situ ICMEs that were not predicted (defined by no arrival around ±1.5 days of the *icme_start_time*). True negatives are more difficult to assess. *Tucker-Hood et al.* [2015] counted the days where no ICME arrived and none was predicted. Here we take a slightly different approach: we take into account the full length of ±1.5 days as T_{window} = 3 days because it defined TP, FP, and FN before. Thus, we count every T_{window} without an ICME nor a predicted arrival as one correct rejection event, and arrive at the total number of TNs by calculating

$$\mathrm{TN} = \left(N_{\mathrm{days}}/T_{\mathrm{window}}\right) - \mathrm{TP} - \mathrm{FP} - \mathrm{FN}, \quad (1)$$

with N_{days} the total number of observation days (2697 days, HIA; 2736 days, HIB). This results in 551/574 (HIA/HIB) intervals of length T_{window} where no CME arrived and none was predicted.

Based on the contingency table entries, several verification measures can be defined [see, e.g., *Woodcock*, 1976; *Jolliffe and Stephenson*, 2006]. Here we focus on the true positive rate (TPR), the false negative rate (FNR), the positive predictive value (PPV), the false alarm ratio (FAR), the threat score (TS), and the bias (BS). The TPR is defined as TP/(TP + FN), the ratio of hits and the total number of observed ICMEs. The FNR is FN/(TP + FN), the ratio of false rejections and the total number of observed ICMEs. The PPV is TP/(TP + FP), the ratio of hits to the total number of predicted ICMEs; and the FAR is FP/(TP + FP), the ratio of false alarms and the total number of predicted events.

In addition, the threat score (TS) is given by TP/(TP + FP + FN), the ratio of hits to the number of all events. The TS is defined in the range [0, 1], where the worst TS is 0, and the best possible TS is 1. The bias (BS) is (TP + FP)/(TP + FN), the ratio of predicted and observed events. The BS indicates that if the forecasts are unbiased (BS = 1), they tend to underforecast (BS < 1) or overforecast (BS > 1) the number of ICMEs. All these measures use parts of the entries in Table 2, whereas the Heidke skill score (HSS) and the Hanssen and Kuipers discriminant, also called True skill statistic (TSS), are calculated from the complete contingency table. The TSS

Table 3. Skill Scores for Earth/L1 for Predictions With the SSEF30 Technique With Heliospheric Imagers and Verified With the Wind Interplanetary CME (ICME) List[a]

Skill Score	Short Form	Results HIA	Results HIB
True positive rate	TPR	0.44	0.43
False negative rate	FNR	0.56	0.57
Positive predictive value	PPV	0.30	0.31
False alarm ratio	FAR	0.70	0.69
Threat score	TS	0.22	0.22
Bias	BS	1.44	1.41
Heidke skill score	HSS	0.17	0.18
True skill statistic	TSS	0.20	0.21

[a]Valid for a time window of ±1.5 days. Total number of CME events: 697 (HIA), 653 (HIB); total predicted arrivals: 250 (HIA), 241 (HIB); total observed ICMEs at Earth/L1: 165.

and HSS are a measure of the overall forecast performance (for definition, see *Bloomfield et al.* [2012] or *Jolliffe and Stephenson* [2006]). We repeat here only the formula for the TSS:

$$\mathrm{TSS} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) - \mathrm{FP}/(\mathrm{FP} + \mathrm{TN}) \quad (2)$$

The TSS is defined in the range [−1, 1], where the worst TSS is 0 and the best possible TSS is 1 or −1 (inverse classification). In contrast, the HSS is defined in the range [−∞, 1], where the worst HSS is smaller than 0 (HSS = 0, indicates no forecast skill) and the best HSS is 1.

In Table 3, we present the skill scores computed for Earth/L1. We find that all verification measures computed are similar for HIA and HIB, respectively. For HIA (HIB) we find that the TPR is 0.44 (0.43), the FNR is 0.56 (0.57), the PPV is 0.30 (0.31), and the FAR is 0.70 (0.69). This means that 44% (43%) of ICMEs are correctly predicted with the SSEF30 technique, while only 30% (31%) of all ICMEs forecasted are observed. The computed BS of 1.44 (1.41) also confirms that the total number of ICMEs is clearly overforecasted. Focusing on the TS we find that 22% (22%) of ICME forecasts among all ICME events (either forecasted or observed) were correct. The high numbers of false alarms and misses are also reflected in the computed HSS and TSS, where the HSS is 0.17 (0.18) and the TSS is 0.20 (0.21).

Regarding the dependence on the length of the time windows, it is very interesting to note that the TSS is independent of the time window between ±1.0 days and ±2.0 days, resulting in very little variation of TSS within 0.19 to 0.22. Outside of this range of time windows, the TSS declines. Considering the differences in the setup used (HI method, ICME list, and time window) our results are comparable with *Tucker-Hood et al.* [2015] who report a TPR of 0.31, a FAR of 0.64, and a HSS of 0.27.

The TSS for STEREO A self-predictions (i.e., HIA verified by STEREO A in situ observations) is only 0.05, because of an extremely high FNR of 0.90. This means that most CMEs that are detected by the in situ instruments are missed by its heliospheric imager. For STEREO B most skill scores and in particular the FNR are comparable to STEREO A, though TSS is here slightly better with a TSS of 0.15.

3.3. Arrival Times, Speeds, and Magnetic Fields

Figure 5 demonstrates the arrival time differences for each spacecraft and HIA/HIB separately (a–e and f–j, respectively). We show the calculated (C) minus observed (O) differences in arrival times in hours. Positive values stand for cases where the CME arrived earlier than the forecast implied, and negative values signify late CME arrivals. Again we use a time window of ±1.5 days around all predicted arrival times in ARRCAT. The choice of 1.5 days is somewhat arbitrary, so we also quote results for ±1.0 and ±2.0 days. The ICME start time that is closest to the predicted arrival is taken for the C-O calculation, and this time difference must be inside the time window for which the C-O is calculated. All other ICMEs inside and outside this time window are ignored. These "double hits," where a predicted CME impact could be related to two ICME start times, are quite seldom and happen, e.g., at Earth only for 12% (9%) of arrivals predicted using HIA (HIB).

The total number of comparisons for HIA was 171, for HIB 194. The average C-O is 2.4 ± 17.1 h for HIA, and 2.7 ± 16.0 h for HIB, so no significant difference. For both spacecraft the average is 2.6 ± 16.6 h. For Earth/L1 alone, this result is 3.2 ± 16.3 h (HIA) and 0.8 ± 17.4 h (HIB). With a time window of 1.0 days, this result is 2.3 ± 12.5 h (HIA), 0.0 ± 12.1 h (HIB), and for 2.0 days its 5.3 ±22.5 h (HIA), and 1.3 ±22.9 h (HIB). This means that an arrival time difference depends on the time window, and thus, a time window should always be given when discussing arrival time differences.

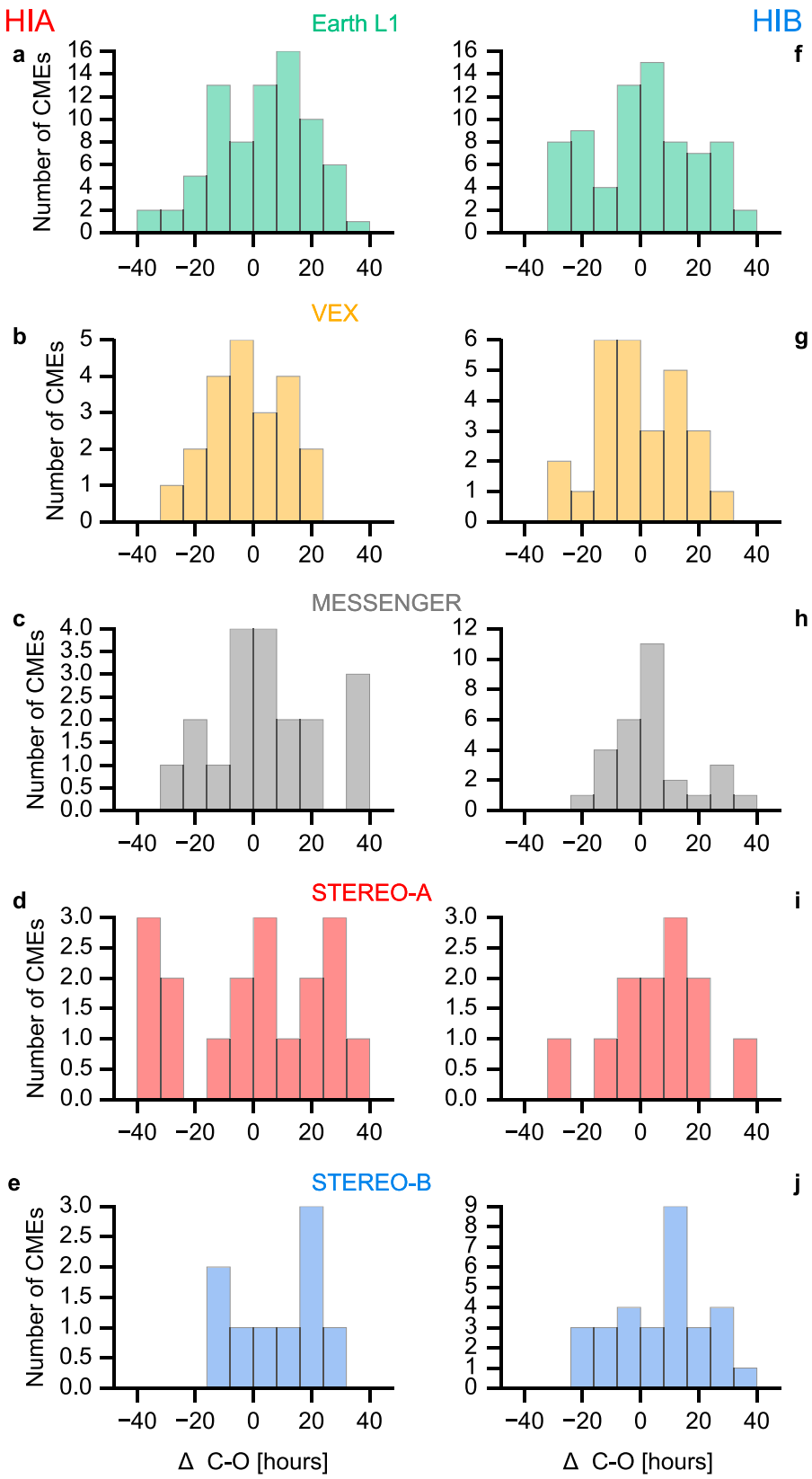

Figure 5. Histograms of calculated (C) minus observed (O) differences in arrival times, in hours for each spacecraft. From top to bottom: Wind, VEX, MESSENGER, STEREO A, STEREO B. Observations based on (a–e) HIA and (f–j) HIB.

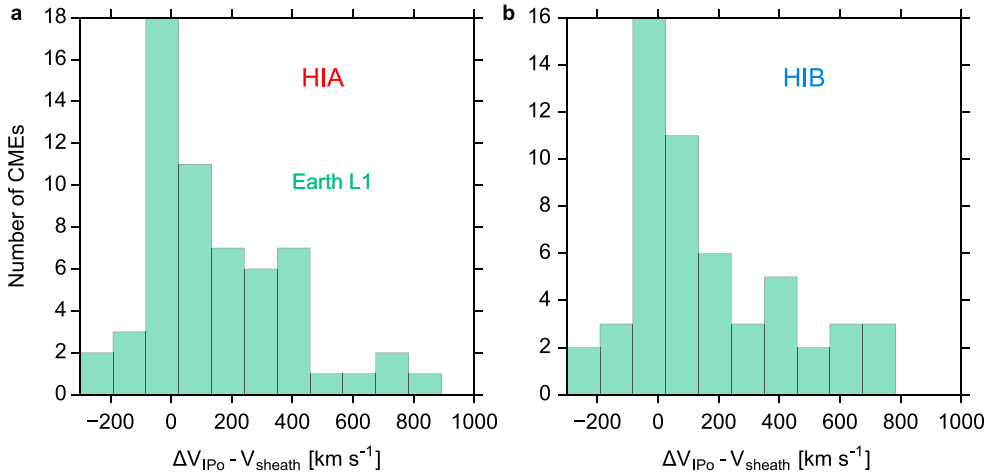

Figure 6. Difference of HI predicted speeds to ICME plasma speeds. (a) Calculated minus observed CME speed at Earth, for HIA. The HI speed is corrected for the circular SSEF30 front. The in situ speed is the proton speed in the ICME sheath region. (b) Same for HIB.

In comparison, *Möstl et al*. [2014] find an average C-O = −1.4 ± 11.1 (their Table 3) for the SSEF30 technique, with all in situ spacecraft at 1 AU. The difference is explained by the selection of events by these authors that quite clearly matched between HI and in situ, whereas here there is no such selection. Most values in Table 2 are positive, so the CME arrived slightly earlier than predicted.

As an experiment, we have excluded events that propagate too far away from the solar equatorial plane in latitude so the only CMEs with a central position angle (PA) between ±20 degree around a PA of 90° (HIA) or 270° (HIB) are included as impacts. This works slightly better concerning the hit and miss percentages with 27% (HIA, a plus of 4%) and 35 (HIB, plus 5%). The values for the arrival time differences are overall here 2.8 ± 15.5 h (HIA) and 2.6 ± 15.2 h (HIB), with no significant change in the averages and improvements by 1.6 (HIA) and 0.8 (HIB) hours in the standard deviations.

Figure 6 demonstrates the calculated minus observed CME speed at Earth. The in situ speed for comparison is the proton speed in the ICME sheath region. It is seen that many events cluster around a speed difference of 0, but there is a positive tail to the distribution, which is a consequence of the constant speed assumption of SSEF30, overestimating the arrival speed [*Möstl et al*., 2014]. The HI speed is corrected for the circular SSEF30 front [*Möstl and Davies*, 2013]. The in situ speed for comparison is the proton speed in the ICME sheath region. The differences are for HIA at Earth L1: 191 ± 341 km s^{-1} and HIB at Earth L1: 245 ± 446 km s^{-1}. These results are very similar to *Möstl et al*. [2014] who quote 252 ± 302 km s^{-1}. Because we do not have plasma speeds from VEX and MESSENGER, we cannot do this comparison for <1 AU.

It is also of great interest to find constraints on the magnitude of in situ ICME magnetic fields from the HI data. Figure 7 visualizes two relationships between parameters derived from HI as part of the arrival catalog and the in situ mean magnetic field in the magnetic obstacle.

Figure 7a shows the predicted speed at the target and the mean magnetic field in the magnetic obstacle, between *mo_start_time* and *mo_end_time*. This is plotted for the hits (true positives) in the data set. Most events cluster around slow to intermediate speeds ($<$1000 km s^{-1}) and low magnetic fields ($<$20 nT). Only at MESSENGER, because of its closer radial distance to the Sun, high average magnetic fields of up to about 90 nT are seen; at all other spacecraft they have been quite low during this solar cycle. For the events classified as hits, the averages ± standard deviations in situ magnetic fields in the magnetic obstacles are for Earth/L1 10.5 ± 5.0 nT, for VEX at Venus 19.8 ± 9.1 nT, and for MESSENGER 43.3 ± 17.9 nT. For a few events where the HI in situ connection was quite clear and only in situ events near 1 AU were used, *Möstl et al*. [2014] derived a linear relationship for a similar plot. In Figure 7b, the predicted travel time, which is defined as *target_arrival-sse_launch*, is plotted against the mean magnetic field. This highlights that only short travel times are accompanied by strong magnetic fields. This relationship does depend on radial distance, as the events from MESSENGER again dominate the left side of the plot.

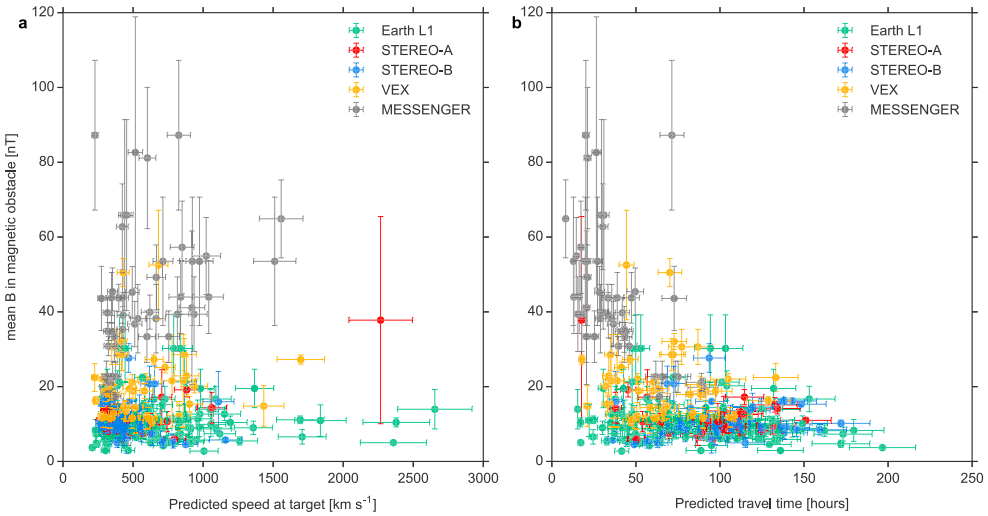

Figure 7. Relationship between predicted speed, travel time and the mean magnetic field in the magnetic obstacle (MO), for all events classified as hits (true positives). (a) The predicted target speed is shown with a typical error bar of ±10% [*Möstl et al*., 2014] against the mean MO field, with the error bar given by the standard deviation of the MO field. (b) The predicted travel time, also with a ±10% error, against the mean MO field. The color code for each spacecraft is given in the plot legend.

These plots and values may give some rough indication of the magnitude of the magnetic field that is to be expected when looking at the predicted CME travel time, the predicted impact speed, and its heliospheric location.

4. Conclusions

We have compared the predictions of the impact and the arrival time from the SSEF30 technique using STEREO/HI observations for more than 1300 CME events with more than 600 ICMEs that were reported in the existing literature and online data sources and merged by us into a singular catalog. Thereby we have derived some fundamental results concerning heliospheric imager-based geometrical modeling with the SSEF30 technique. This is of major importance for a space weather mission carrying a heliospheric imager at the L5 or L1 point *[Lavraud et al*., 2016; *DeForest et al*. 2016], and all other missions that carry heliospheric imagers (*STEREO*, *Solar Orbiter*, and *Solar Probe Plus*). In summary, our results derived from science data and those by *Tucker-Hood et al*. [2015] for HI real time data show that HI modeling is currently at the same level for predicting CME arrivals as are other analytical and numerical techniques.

4.1. Contingency Table Results

We find that one out of four predicted impacts with the SSEF30 technique results in a clear ICME that includes a magnetic obstacle, for a time window of ±1.5 days around the predicted arrival time. Including a constraint on the propagation latitude of the CME in HI, so that only CMEs which centrally propagate near the solar equatorial plane are selected, improves the hit percentages to about 33%, so two false alarms for one correct prediction.

We find for time windows from 1.0 to 2.0 days a stable true skill statistic (TSS) of 0.19 to 0.22, similar for HIA and HIB. What does this mean in comparison to other studies? We need to caution that we have used science data to discuss hindsight predictions of CMEs with heliospheric imagers. *Tucker-Hood et al*. [2015], analyzing the real time performance of HI with beacon data, find a slightly higher TSS of 0.26 (calculated from their contingency table). The higher TSS than in our study is likely explained by smaller statistics and a longer time window as in our study. For several techniques to calculate the shock arrivals, *Zhao and Dryer* [2014] find TSS between 0.15 and 0.25, similar to our results.

Overall, the TSS of about 0.20 might be considered rather low concerning the accuracy of a regular space weather forecast. However, we recognize several aspects that are expected to improve the prediction accuracy. First, the HI SSEF30 technique does not adequately reflect basic CME physics (see below). Second, our analysis relied on existing ICME lists, focusing on clear ICMEs that feature magnetic flux ropes. We did not include ICME flank encounters that often feature unclear ICME signatures [e.g., *Richardson and Cane*, 2010]

and sometimes "driverless shocks" [e.g., *Gopalswamy et al*., 2010; *Janvier et al*., 2014], i.e., the whole bulk of the CME is missed. Even such cases are interesting for space weather as CME sheaths alone can drive major storms [*Huttunen et al*., 2002; *Huttunen and Koskinen*, 2004].

In future work, the forecasted CME arrivals might be checked against the in situ data directly, not just the ICME lists, which would eliminate the identification of ICMEs which is always a subjective matter. Another clear improvement is to take into account event specific aspects. A major assumption that needs to be eliminated is the constant 60° full width of the CME (based on *Yashiro et al*. [2004] and determine the width for each event separately. This emphasizes also the importance of multispacecraft missions to allow CME reconstructions from multiple vantage points to obtain realistic parameter that are not subject to projection effects [e.g., *Thernisien et al*., 2009].

4.2. Arrival Times

Forecasting with HI also works well for locations at <1 AU, as we find that the predicted arrival times match with similar accuracy compared for spacecraft positioned around 1 AU. This result might be expected for Mars too, though we could not yet test this as MAVEN entered its orbit around Mars in late 2014 just when STEREO went into conjunction. The mean, signed arrival time differences of around $\pm$17 h are very close to those reported by a comparison of the heliospheric MHD simulation Enlil and the drag-based model forecasts at Earth [*Vršnak et al*., 2014]. Our sample for this comparison consists of 315 CMEs, 143 observed by HIA, and 172 by HIB. *Mays et al*. [2015] report a mean absolute difference of 12.3 h for predicted shock arrivals with Enlil at Earth; *Zhao and Dryer* [2014] quote 10 h. For our study and looking only at the Earth/L1 arrivals, this mean absolute difference is only slightly larger: 14.0 (HIA) and 14.4 h (HIB), based on 76 (HIA) and 74 (HIB) comparisons to in situ data. However, it must be noted that such comparisons depend on the choice of the time window for successful hits. In summary, with the current SSEF30 technique, using an HI instrument does not enhance the CME prediction accuracy, but its results are perfectly comparable with those from other methods that use analytical and numerical modeling.

To improve the arrival time predictions, in future work, the SSEF30 modeling results on all 1337 CMEs should be updated with ElEvoHI [*Rollett et al*., 2016], as the constant speed of SSEF30 assumption runs into troubles in particular predicting the arrival speeds. A robust automatization of ElEvoHI is needed to do this. ElEvoHI eliminates both assumptions of a circular front and constant speed and instead uses an elliptical shape [*Janvier et al*., 2014; *Möstl et al*., 2015; *Rollett et al*., 2016] and a decelerating speed profile [*Vršnak et al*., 2013]. The ellipse shape may even evolve with time. We expect that applying ElEvoHI to all CME tracks in the HIGeoCat will show improvements concerning the arrival time and arrival speed. Future work should also focus on adding magnetic fields to predictions with heliospheric imagers, by bringing together a model such as FRi3D [*Isavnin*, 2016] or FIDO [*Kay et al*., 2017] with ElEvoHI to make predictions of CME geomagnetic effects possible [*Tobiska et al*., 2013; *Kubicka et al*., 2016].

4.3. Implications for Future Missions

We have demonstrated that L5 could be a good location for the prediction of Earth-directed CMEs with HI. But somewhat surprisingly, if the HI observing spacecraft is positioned further away from the Earth in heliospheric longitude, the hit/miss predictions do not get worse. Note, however, that this was already seen by *Lugaz et al*. [2012] and *Möstl et al*. [2014]. We can now base this conclusion on a larger event sample, but there are other factors such as a changing solar activity from minimum to maximum that may considerably influence this result, and more specific work should be dedicated to understanding this effect. In any case, this has an implication for current operations with STEREO Ahead: real-time predictions with HI should be done even when the spacecraft is still behind the east limb of the Sun as seen from Earth. This means that we do not have to wait to use HI until it passes L5 in July 2020 to make good predictions using the HI instrument.

It has also been proposed that L1 could be a good location for a heliospheric imager [*DeForest and Howard*, 2015]. While we find that "self-predictions" of CMEs observed with HIA or HIB and later detected with the same spacecraft by the in situ instruments measuring magnetic fields and plasma parameters show slightly better percentages of correct hits than those to other spacecraft, the rate of false rejections is much higher than for an "away" view point, meaning most ICMEs that impacted STEREO A or B were not detected by its own heliospheric imager.

5. Sources of Data and Supplementary Material

Catalogs:

HIGeoCat: https://www.helcats-fp7.eu/catalogues/wp3_cat.html
ARRCAT: doi:10.6084/m9.figshare.4588324
https://doi.org/10.6084/m9.figshare.4588324.v1
ICMECAT: doi:10.6084/m9.figshare.4588315
https://doi.org/10.6084/m9.figshare.4588315.v1

ICME lists:

Wind ICME list: T. Nieves-Chinchilla et al. https://wind.nasa.gov/ICMEindex.php
STEREO ICME list: Lan Jian, http://www-ssc.igpp.ucla.edu/forms/stereo/stereo_level_3.html updated by our study.
VEX: *Good and Forsyth* [2016], updated by our study.
MESSENGER: *Winslow et al.* [2015], *Good and Forsyth* [2016], updated by our study.

In situ data:

Wind: https://cdaweb.sci.gsfc.nasa.gov
STEREO: http://aten.igpp.ucla.edu/forms/stereo/level2_plasma_and_magnetic_field.html
MESSENGER: https://pds-ppi.igpp.ucla.edu
VEX: data obtained from magnetometer PI T.L. Zhang tielong.zhang@oeaw.ac.at

Animation of Figure 2:

doi.org/10.6084/m9.figshare.4602253
https://www.youtube.com/watch?v=Jr4XRzGCaaQ&

Acknowledgments
The presented work has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement 606692 (HELCATS). This study was supported by the Austrian Science Fund (FWF): P26174-N27. L.P. acknowledges support from the Natural Sciences and Engineering Research Council (NSERC) of Canada. V.K. acknowledges the support of the Czech Science Foundation grant 17-06818Y. "R.M.W. is supported by NASA grant NNX15AW31G and by NSF grant AGS1622352." This research has made use of SunPy, an open-source and free community-developed solar data analysis package written in Python [*The SunPy Community*, 2015]. We would like to thank the Editor and two reviewers for helpful suggestions to improve the manuscript. All sources of data that were used in producing the results presented in this study are quoted in section 5.

References

Anderson, B. J., M. H. Acuna, D. A. Lohr, J. Scheifele, A. Raval, H. Korth, and J. A. Slavin (2007), The magnetometer instrument on MESSENGER, *Space Sci. Rev.*, *131*, 417, doi:10.1007/s11214-007-9246-7.

Bloomfield, D. S., P. A. Higgins, R. T. J. McAteer, and P. T. Gallagher (2012), Toward reliable benchmarking of solar flare forecasting methods, *Astrophys. J.*, *747*, L41, doi:10.1088/2041-8205/747/2/L41.

Colaninno, R. C., A. Vourlidas, and C. C. Wu (2013), Quantitative comparison of methods for predicting the arrival of coronal mass ejections at Earth based on multiview imaging, *J. Geophys. Res. Space Physics*, *118*, 6866–6879, doi:10.1002/2013JA019205.

Davies, J. A., et al. (2012), A self-similar expansion model for use in solar wind transient propagation studies, *Astrophys. J.*, *750*, 23, doi:10.1088/0004-637X/750/1/23.

Davis, C. J., J. A. Davies, M. Lockwood, A. P. Rouillard, C. J. Eyles, and R. A. Harrison (2009), Stereoscopic imaging of an Earth-impacting solar coronal mass ejection: A major milestone for the STEREO mission, *Geophys. Res. Lett.*, *36*, L08102, doi:10.1029/2009GL038021.

DeForest, C. E., and T. A. Howard (2015), Feasibility of heliospheric imaging from near Earth, *Astrophys. J.*, *804*, 126, doi:10.1088/0004-637X/804/2/126.

DeForest, C. E., T. A. Howard, D. F. Webb, and J. A. Davies (2016), The utility of polarized heliospheric imaging for space weather monitoring, *Space Weather*, *14*, 32–49, doi:10.1002/2015SW001286.

Eyles, C. J., et al. (2009), The heliospheric imagers onboard the STEREO mission, *Sol. Phys.*, *254*, 387–445, doi:10.1007/s11207-008-9299-0.

Galvin, A. B., et al. (2008), The Plasma and Suprathermal Ion Composition (PLASTIC) investigation on the STEREO observatories, *Space Sci. Rev.*, *136*, 437–486, doi:10.1007/s11214-007-9296-x.

Good, S. W., and R. J. Forsyth (2016), Interplanetary coronal mass ejections observed by MESSENGER and Venus Express, *Sol. Phys.*, *291*, 239, doi:10.1007/s11207-015-0828-3.

Gopalswamy, N., P. Mäkelä, H. Xie, S. Akiyama, and S. Yashiro (2010), Solar sources of "driverless" interplanetary shocks, in *Twelfth International Solar Wind Conference*, vol. 2126, pp. 452–458, Am. Inst. Phys. Melville, New York, doi:10.1063/1.3395902.

Howard, R. A., et al. (2008), Sun Earth connection coronal and heliospheric investigation (SECCHI), *Space Sci. Rev.*, *136*, 67–115, doi:10.1007/s11214-008-9341-4.

Huttunen, K. E J., H. E. J. Koskinen, T. I. Pulkkinen, A. Pulkkinen, and M. Palmroth (2002), April 2000 magnetic storm: Solar wind driver and magnetospheric response, *J. Geophys. Res.*, *107*(A12), 1440, doi:10.1029/2001JA009154.

Huttunen, K. E. J., and H. E. J. Koskinen (2004), Importance of post-shock streams and sheath regions as driver of intense magnetospheric storms and high latitude activity, *Ann. Geophys.*, *22*, 1729–1738, doi:10.5194/angeo-22-1729-2004.

Isavnin, A. (2016), FRiED: A novel three-dimensional model of coronal mass ejections, *Astrophys. J.*, *833*, 267, doi:10.3847/1538-4357/833/2/267.

Janvier, M., P. Démoulin, and S. Dasso (2014), Mean shape of interplanetary shocks deduced from in situ observations and its relation with interplanetary CMEs, *Astron. Astrophys.*, *565*, A99, doi:10.1051/0004-6361/201423450.

Jian, L. K., C. T. Russell, J. G. Luhmann, and R. M. Skoug (2006), Properties of interplanetary coronal mass ejections at one AU during 1995–2004, *Sol. Phys.*, *239*, 393, doi:10.1007/s11207-006-0133-2.

Jian, L. K., C. T. Russell, J. G. Luhmann, A. B. Galvin, and K. D. C. Simunac (2013), Solar wind observations at STEREO: 2007–2011, in *Proceedings of Solar Wind 13*, *Am. Inst. Phys.*, vol. 1539, pp. 191–194, Melville, New York, doi:10.1063/1.4811020.

Jolliffe, I. T., and D. B. Stephenson (2006), *Forecast Verification. A Practitioner's Guide in Atmospheric Science*, 2nd ed., Wiley, Hoboken, N. J.

Kay, C., N. Gopalswamy, A. Reinard, and M. Opher (2017), Predicting the magnetic field of Earth-impacting CMEs, *Astrophys. J.*, *835*, 117, doi:10.3847/1538-4357/835/2/117.
Kilpua, E. K. J., A. Isavnin, A. Vourlidas, H. E. J. Kosinen, and L. Rodriguez (2013), On the relationship between coronal mass ejections and magnetic clouds, *Ann. Geophys.*, *31*, 1251, doi:10.5194/angeo-31-1251-2013.
Kubicka, M., C. Möstl, T. Amerstorfer, P. D. Boakes, L. Feng, J. P. Eastwood, and O. Törmänen (2016), Prediction of geomagnetic storm strength from inner heliospheric in situ observations, *Astrophys. J.*, *833*, 255, doi:10.3847/1538-4357/833/2/255.
Lavraud, B., et al. (2016), A small mission concept to the Sun-Earth Lagrangian L5 point for innovative solar, heliospheric and space weather science, *J. Atmos. Sol. Terr. Phys.*, *146*, 171, doi:10.1016/j.jastp.2016.06.004.
Lepping, R., et al. (1995), The WIND magnetic field investigation, *Space Sci. Rev.*, *71*, 207, doi:10.1007/BF00751330.
Liu, Y., A. Thernisien, J. G. Luhmann, A. Vourlidas, J. A. Davies, R. P. Lin, and S. D. Bale (2010), Reconstructing coronal mass ejections with coordinated imaging and in situ observations: Global structure, kinematics, and implications for space weather forecasting, *Astrophys. J.*, *722*, 1762, doi:10.1088/0004-637X/722/2/1762.
Lugaz, N. (2010), Accuracy and limitations of fitting and stereoscopic methods to determine the direction of coronal mass ejections from heliospheric imagers observations, *Sol. Phys.*, *267*, 411, doi:10.1007/s11207-010-9654-9.
Lugaz, N., P. Kintner, C. Möstl, L. K. Jian, C. J. Davis, and C. J. Farrugia (2012), Heliospheric observations of STEREO-directed coronal mass ejections in 2008–2010: Lessons for future observations of Earth-directed CMEs, *Sol. Phys.*, *279*, 497, doi:10.1007/s11207-012-0007-8.
Luhmann, J. G., et al. (2008), STEREO IMPACT investigation goals, measurements, and data products overview, *Space Sci. Rev.*, *136*, 117, doi:10.1007/s11214-007-9170-x.
Mays, M. L., et al. (2015), Ensemble modeling of CMEs using the WSA–ENLIL+Cone model, *Sol. Phys.*, *290*, 1775, doi:10.1007/s11207-015-0692-1.
Murray, S. A., S. Bingham, M. Sharpe, and D. R. Jackson (2017), Flare forecasting at the Met Office Space Weather Operations Centre, ArXiv e-prints: https://arxiv.org/abs/1703.06754.
Möstl, C., and J. A. Davies (2013), Speeds and arrival times of solar transients approximated by self-similar expanding circular fronts, *Sol. Phys.*, *285*, 411–423.
Möstl, C., M. Temmer, T. Rollett, C. J. Farrugia, Y. Liu, A. M. Veronig, M. Leitner, A. B. Galvin, and H. K. Biernat (2010), STEREO and Wind observations of a fast ICME flank triggering a prolonged geomagnetic storm on 5-7 April 2010, *Geophys. Res. Lett.*, *37*, L24103, doi:10.1029/2010GL045175.
Möstl, C., et al. (2014), Connecting speeds, directions and arrival times of 22 coronal mass ejections from the Sun to 1 AU, *Astrophys. J.*, *787*, 119.
Möstl, C., et al. (2015), Strong coronal channelling and interplanetary evolution of a solar storm up to Earth and Mars, *Nat. Commun.*, *6*, 7135, doi:10.1038/ncomms8135.
Ogilvie, K. W., et al. (1995), A comprehensive plasma instrument for the wind spacecraft, *Space Sci. Rev.*, *71*, 55, doi:10.1007/BF00751326.
Oughton, E. J., A. Skelton, R. B. Horne, A. W. P. Thomson, and C. T. Gaunt (2017), Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure, *Space Weather*, *15*, 65–83, doi:10.1002/2016SW001491.
Reiss, M. A., M. Temmer, A. M. Veronig, L. Nikolic, S. Vennerstrom, F. Schöngassner, and S. J. Hofmeister (2016), Verification of high-speed solar wind stream forecasts using operational solar wind models, *Space Weather*, *14*, 495–510, doi:10.1002/2016SW001390.
Richardson, I. G., and H. V. Cane (2010), Near-Earth interplanetary coronal mass ejections during solar cycle 23 (1996–2009): Catalog and summary of properties, *Sol. Phys.*, *264*, 189–237, doi:10.1007/s11207-010-9568-6.
Rollett, T., C. Möstl, A. Isavnin, J. A. Davies, M. Kubicka, U. V. Amerstorfer, and R. A. Harrison (2016), ElEvoHI: A novel CME prediction tool for heliospheric imaging combining an elliptical front with drag-based model fitting, *Astrophys. J.*, *824*, 131, doi:10.3847/0004-637X/824/2/131.
Rouillard, A. P. (2011), Relating white light and in situ observations of coronal mass ejections: A review, *J. Atmos. Sol. Terr. Phys.*, *73*, 1201, doi:10.1016/j.jastp.2010.08.015.
Rouillard, A. P., et al. (2008), First imaging of corotating interaction regions using the STEREO spacecraft, *Geophys. Res. Lett.*, *35*, L10110, doi:10.1029/2008GL033767.
Sheeley, N. R., J. H. Walters, Y.-M. Wang, and R. A. Howard (1999), Continuous tracking of coronal outflows: Two kinds of coronal mass ejections, *Geophys. Res. Lett.*, *104*, 24,739–24,767, doi:10.1029/1999JA900308.
Thernisien, A., A. Vourlidas, and R. A. Howard (2009), Forward Modeling of Coronal Mass Ejections Using STEREO/SECCHI Data, *Sol. Phys.*, *256*, 111, doi:10.1007/s11207-009-9346-5.
The SunPy Community, et al. (2015), SunPy—Python for solar physics, *Comput. Sci. Discovery*, *8*, 014009, doi:10.1088/1749-4699/8/1/014009.
Thompson, W. T. (2006), Coordinate systems for solar image data, *Astron. Astrophys.*, *449*, 791, doi:10.1051/0004-6361:20054262.
Tobiska, W. K., D. Knipp, W. J. Burke, D. Bouwer, J. Bailey, D. Odstrcil, M. P. Hagan, J. Gannon, and B. R. Bowman (2013), The *Anemomilos* prediction methodology for *Dst*, *Space Weather*, *11*, 490–508, doi:10.1002/swe.20094.
Tucker-Hood, K., et al. (2015), Validation of a priori CME arrival predictions made using real-time heliospheric imager observations, *Space Weather*, *13*, 35, doi:10.1002/2014SW001106.
Vršnak, B., et al. (2013), Propagation of interplanetary coronal mass ejections: The drag-based model, *Sol. Phys.*, *285*, 295, doi:10.1007/s11207-012-0035-4.
Vršnak, B., M. Temmer, T. Žic, A. Taktakishvili, M. Dumbović, C. Möstl, A. M. Veronig, M. L. Mays, and D. Odstrčil (2014), Heliospheric propagation of coronal mass ejections: Comparison of numerical WSA-ENLIL+cone model and analytical drag-based model, *Astrophys. J. Suppl. Ser.*, *213*, 21, doi:10.1088/0067-0049/213/2/21.
Winslow, R. M., B. J. Anderson, C. L. Johnson, J. A. Slavin, H. Korth, M. E. Purucker, D. N. Baker, and S. C. Solomon (2013), Mercury's magnetopause and bow shock from MESSENGER Magnetometer observations, *J. Geophys. Res. Space Physics*, *118*, 2213–2227, doi:10.1002/jgra.50237.
Winslow, R. M., N. Lugaz, L. C. Philpott, N. A. Schwadron, C. J. Farrugia, B. J. Anderson, and C. W. Smith (2015), Interplanetary coronal mass ejections from MESSENGER orbital observations at Mercury, *J. Geophys. Res. Space Physics*, *120*, 6101–6118, doi:10.1002/2015JA021200.
Wood, B. E., Wu, C.-C., Lepping, R. P., Nieves-Chinchilla, T., Howard, R. A., Linton, M. G., and D. G. Socker (2017), A STEREO survey of magnetic cloud coronal mass ejections observed at Earth in 2008–2012, ArXiv e-prints: https://arxiv.org/abs/1701.01682.
Woodcock, F. (1976), The evaluation of yes/no forecasts for scientific and administrative purposes, *Mon. Weather Rev.*, *104*, 1209, doi:10.1175/1520-0493(1976)104<1209:TEOYFF>2.0.CO;2.
Yashiro, S., N. Gopalswamy, G. Michalek, O. C. St. Cyr, S. P. Plunkett, N. B. Rich, and R. A. Howard (2004), A catalog of white light coronal mass ejections observed by the SOHO spacecraft, *J. Geophys. Res.*, *109*, A07105, doi:10.1029/2003JA010282.
Zhang, T. L., et al. (2006), Magnetic field investigation of the Venus plasma environment: Expected new results from Venus Express, *Planet. Space Sci.*, *54*, 1336, doi:10.1016/j.pss.2006.04.018.
Zhang, T. L., et al. (2008), Initial Venus Express magnetic field observations of the Venus bow shock location at solar minimum, *Planet. Space Sci.*, *56*, 785, doi:10.1016/j.pss.2007.09.012.
Zhao, X., and M. Dryer (2014), Current status of CME/shock arrival time prediction, *Space Weather*, *12*, 448–469, doi:10.1002/2014SW001060.